

**Figure 1: Top 10 BLASTP hits for INSP179 polypeptide sequence (SEQ ID NO:10)  
against NCBI-nr**

BLASTP 2.2.2 [Jan-08-2002]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= INSP179.pp  
(880 letters)

Database: All non-redundant GenBank CDS  
translations+PDB+SwissProt+PIR+PRF  
1,594,288 sequences; 522,190,286 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
ref XP_291099.2  similar to hypothetical protein MGC38937 [Homo ...	1747	0.0
ref NP_705796.1  hypothetical protein MGC38937 [Mus musculus] >g...	684	0.0
ref XP_223356.1  similar to hypothetical protein MGC38937 [Rattu...	650	0.0
ref XP_221358.2  similar to Carboxypeptidase N 83 kDa chain (Car...	82	5e-14
ref NP_570843.1  leucine rich repeat containing 15; leucine-rich...	77	2e-12
gb AAH25836.1  1300018K11Rik protein [Mus musculus]	75	7e-12
ref XP_148373.1  RIKEN cDNA 1300018K11 [Mus musculus] >gi 380806...	75	7e-12
ref NP_659551.1  leucine-rich repeat protein induced by beta-amy...	74	1e-11
dbj BAD01045.1  toll-like receptor 3 [Paralichthys olivaceus]	73	3e-11
dbj BAB85498.1  18 wheeler [Bombyx mori]	72	5e-11

**Figure 2: Pairwise alignment of INSP179 to top annotated BLASTP hit (mouse ortholog)**

```

>ref|NP_705796.1| hypothetical protein MGC38937 [Mus musculus]
gb|AAH31901.1| Hypothetical protein MGC38937 [Mus musculus]
dbj|BAC34797.1| unnamed protein product [Mus musculus]
      Length = 872

Score = 684 bits (1765), Expect = 0.0
Identities = 422/888 (47%), Positives = 529/888 (59%), Gaps = 41/888 (4%)

Query: 1  MKNLYFRVITIVIGLYFTGIMTNASRKSNIENSECQWNEYILTNCSTGKCDIPVDISQ 60
      M++ Y RV +V GL F +T SRKS++ FN E Q N +L N S +SQ
Sbjct: 1  MRDFYVRVTILVTGLCFVETVTPSRKSSVSFNPEYQNRGDLNVNWSIRH-----VSQ 54

Query: 61  TAATVDVSFNFRRVLLQSHTKKEEWKIKHLDLSNNLISKITLSPFAYLHALEVLNLSNNA 120
      +D SF FFRVL Q HT+KE IK D +++ ISK+TL P A+LHALE+LNLSN A
Sbjct: 55  NTDAMDRSFYFFRVLFQPHQKER-HIKPPDRTHHRISKVTLDPPLAHLHALEILNLSNKA 113

Query: 121  IHSLSLDLLSPKSSWVKRHRSSFRNRFLLKVLILQRNKLSDTPKGLWKLKSLQSLDLSF 180
      IH SLD P SS KRH +R P L+VLILQRN+LS TPKGLWKLKSL+SLDLSF
Sbjct: 114  IHYFSLDQPLPSSSHQKRHGGHSHSRPLRQLVILQRNQLSGTPKGLWKLKSLRSLDLSF 173

Query: 181  NGILQIGWSDFHNCLOLENLCLKSNKIFKIPPQAFKDLKKLQVIDLSNNALITILEMMII 240
      N I+ IG SDFH CLQLE++ LKSNKI I P+AFK LKKLQV+DL +NAL T++P++ I
Sbjct: 174  NRIVHIGLSDFHGCLOLESIIYLKSNKICTIHPKAFKGLKKLQVVDLRSNALTTLVPIVTI 233

Query: 241  ALEFPHLVVDLADNNWQCDDSVAVFQNFISESWRKKWNVICNRSIGSEEANGGTPQSRIS 300
      ALE PHL + LADN WQC +S FQN S SWR+ W ICN S+ ++ N T Q R S
Sbjct: 234  ALELPHLELGLADNQWCSESNVNFQNTSSSWREIWKAICNMSVENKRPNAETHQIRKS 293

Query: 301  RETRL---PPIHLHRMKSLIRSKAERPQGGRTGTISTLGKKAKAGSG-LRKKQRRRLPRSV 356
      R+T L PP L KSLI+SKAERPQ G +S LGK+AK G G LR + P +
Sbjct: 294  RDTHLLLSPPSDL---KSLIQSKAERPQAGMDMHLALGKEAKDGYGDLRGMWPQSPVEL 350

Query: 357  RSTRDVQAAGKKEDAPQDLALAVCLSVFITFLVAFSLGAFTRPYVDRLWQKKCQSKSPGL 416
      R ++D Q +K+D P L LA+CLSVFITF+VAF LGAF RPY+DRL Q++C +K PG
Sbjct: 351  RDSQDEQVTDKDDKPPALELAICLSVFITFVVAFCLGAFARPYIDRLRQQRCSNKRPGS 410

Query: 417  DNAYSNEGFYDDMEAGHTPHPHETHLRQVFPHLSLYENQTPFVWTQPHPHATVIPDRTL 476
      DNAYSN+GF+ D+E A H + T L Q HL L ENQ P WV +P PH+ V ++ LG
Sbjct: 411  DNAYSNGKGFHGDIEGAQHMEYQGTDLHQTHHLHLSNQNPSSWVAEPIPHSAVQSEQMLG 470

Query: 477  RSRKDPGSSQSPGQCGDNTGAGSGND-----GAVYSILQRHPHAGNRELMASAAQDHIHR 530
      + DPG QSP Q D+ + SG+ + L P+A + +S Q H
Sbjct: 471  SNGTDPGHQQSPEQLKDSNESRSGDSIVLPSGPVAHLALHGLPNADAHKAISPVQ---HH 527

Query: 531  NDILGEWTYETVAQEEPLSAHSGVGVSSVAGTSHAVSGSSRYDSNELDPSLSGEITASLCK 590
      +D L E Y+TVAQE L + SS+ G S ++L PS ++ AS K
Sbjct: 528  HDFLEEAHYDTVAQEYSLIDDVMDRSSITGPLGTFPSSVESRRDDLHPSQPRDVVASFSK 587

Query: 591  MLTHAEAQRGTGDSKERGGTEQ-SLWDSQMEFSKERQVSSSIDLLSIQQPRLSGARAEAL 649
      L HA + S E G E DSQM S+ERQVS+SI L+ QQP G AEE L
Sbjct: 588  TLAHANTREAEGSMETGCPEPLGAMDSQMGSSSEERQVNSIRELATQQPSFQGVDAEERL 647

Query: 650  SAHYSEVPYGDPRDTGPSVFPPRWDSGLDVTANKEPVQKSTPSTDCCELES--DCDSDE 707
      S YSEV + DP PS+ PRW SG V PA EPV++ P D +L + + DSDE
Sbjct: 648  SHVYSEVLHNDP-----PSL-RPRWGS GHYVIPATGEPVERDAPFDPHYDLVTNYESDSDE 702

Query: 708  GSLFTLSSISSSESARSKTEEAVPDE-----ESLQDESSGASKDNVTAVDSLEENVTFQTI 762
      GSLFTLSS SE RS EE E + L + G KD+VT+ +S+E+ +T Q I
Sbjct: 703  GSLFTLSSEGEDTRSLAEEQASVENDGTSQFLPSRNLGEYKDSVTSAESVED-LTSQRI 761

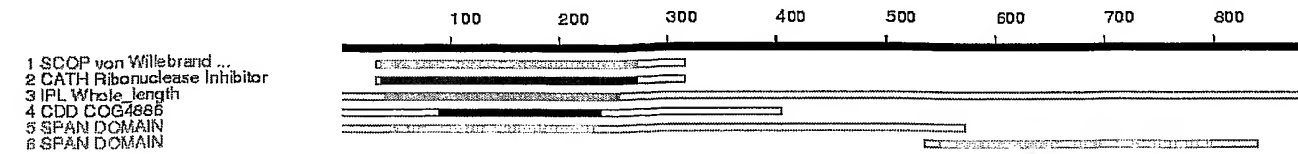
Query: 763  PGKCKNQEDPFEKPLISAPDSGMYKTHLENASDTRSEGLSPWPRSPGNSPLGDEFPGMF 822
      P KC+ QE LIS PDS + +T+ EN S + E S WP+ PG+ E G +
Sbjct: 762  PEKCEAQEAHLRNTLISGPDSCVETNQENDSSSLDPENRSTWPLPGHKLSHHETLGTY 821

```

Query: 823 TYDYDTALQSKAAEWHCSLRDLEFSNVDVLQQTPPCSAEVPSPDPKAA 870  
D QS+A +WH SLRDLE NVD +PP S E S P+ A  
Sbjct: 822 G---DIEPQSEAVDWHYSLRDLESPNVD-SSPSPFYSEDLGSPEDRA 865

**Figure 3: Domain Professor results for INSP179**

Query: INSP179.pp

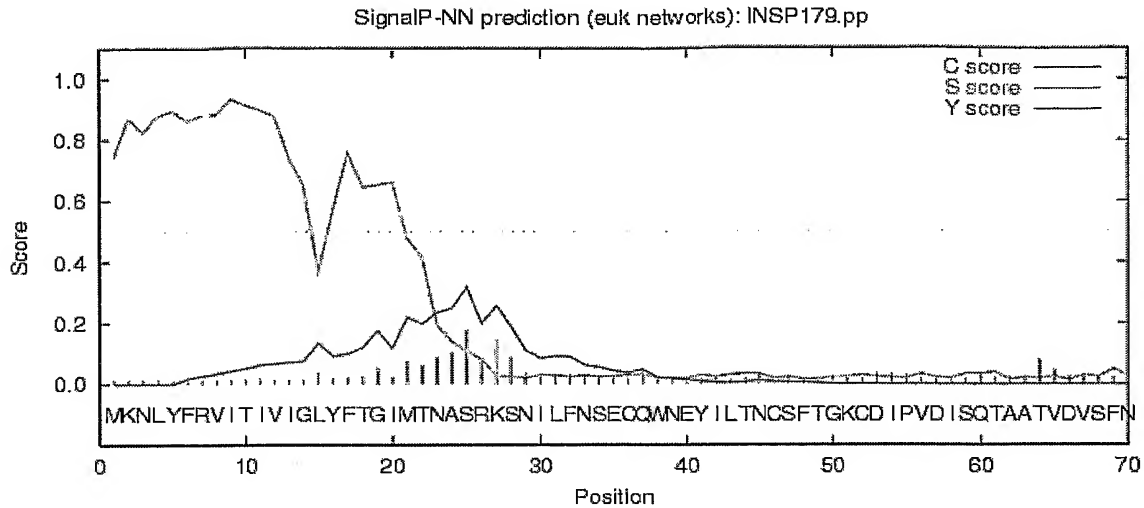


E-values and regions on this page use rep-silon and are approximate - hit "Aln" for Blastpgp  
 e-values and alignments

1	SCOP von Willebrand ...	82%	c.10.2.7 (A:) von Willebrand factor binding domain of glycoprotein Ib alpha {Human (Homo sapiens)}	<a href="#">d1gwba</a>	35-271	4-221	- / -	<u>2</u> / 1.4e-25
2	CATH Ribonuclease Inhibitor	82%	3.80.10.10 Alpha Beta; Horseshoe; Ribonuclease Inhibitor	<a href="#">1gwba0</a>	35-271	4-221	- / -	<u>2</u> / 1.4e-25
3	IPL Whole_length	8%	Ferredoxin-like / Adenyl and guanylyl cyclase catalytic domain Ascomycotal and Fungal ACs, most of the full length alignment. Have added hCP46367.1 and hCP46942.1 to the alignment for this profile.	<a href="#">IPL002890</a>	40-255	1060-1254	<u>1</u> / 1.4e-15	- / -
4	CDD COG4886	33%	Leucine-rich repeat (LRR) protein [Function unknown]	<a href="#">COG4886</a>	90-238	98-228	<u>1</u> / 5.3e-14	- / -
5	SPAN DOMAIN	31%	DOMAIN EXTRACELLULAR (POTENTIAL) sp LGR5_HUMAN O75473 1-561	<a href="#">O75473</a>	45-231	51-222	<u>1</u> / 3.0e-11	- / -
6	SPAN DOMAIN	81%	DOMAIN ASP/SER-RICH sp SR40_YEAST P32583 25-314	<a href="#">P32583</a>	549-797	15-248	<u>1</u> / 9.0e-05	- / -

**Figure 4: Signal peptide prediction (SignalP V2.0) for INSP179 polypeptide sequence (SEQ ID NO: 10).**

>INSP179.pp

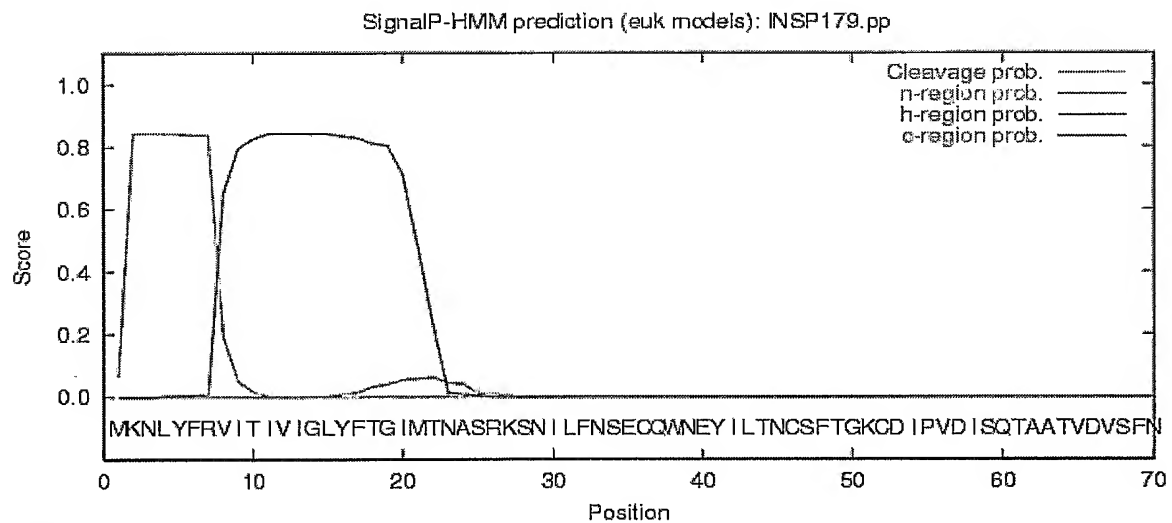


>INSP179.pp

length = 70

# Measure	Position	Value	Cutoff	signal peptide?
max. C	25	0.179	0.33	NO
max. Y	25	0.318	0.32	NO
max. S	9	0.933	0.82	YES
mean S	1-24	0.697	0.47	YES

# Most likely cleavage site between pos. 24 and 25: TNA-SR



>INSP179.pp

Prediction: Signal anchor

Signal peptide probability: 0.065

Signal anchor probability: 0.778

Max cleavage site probability: 0.031 between pos. 24 and 25

**Figure 5: INSP179 DNA and polypeptide sequence**

```

1  ttccgaaatc ttaaattgaa aattaaattt gctgcttatt gctgttagtt tttatatatt
61  attgtgaata agtcaattca ttttcctttt tccaggcatt atgaaaaacc tctatttcag
                                     m k n l y f
                                     INSP179-CP1

121  agtcattacc atagttatag gtctttattt tactggaata atgacaaatg catcaagaaa
    r v i t i v i g l y f t g i m t n a s r
    ──────────▶

181  aagcaatatt ttattcaatt ctgaatgcc aatggaatgaa tatattctga caaattgttc
    k s n i l f n s e c q w n e y i l t n c

241  ttttaccgga aagtgtgata tacctgtgga catatcacag acagcagcca ctgtggatgt
    s f t g k c d i p v d i s q t a a t v d

301  aagtttcaat ttcttttagag ttctcttaca gtctcacacg aaaaaagaag agtggaaaat
    v s f n f f r v l l q s h t k k e e w k

361  aaaacatctg gacctcagta acaatctcat atcaaaaata accttaagcc cttttgcata
    i k h l d l s n n l i s k i t l s p f a

421  tttacatgct ttggaagtgt taaacctcag caacaatgcc atccactccc tctcattgga
    y l h a l e v l n l s n n a i h s l s l

481  tctactcagt cctaagtcct catgggtgaa acgccacaga agcagcttca gaaacagggt
    d l l s p k s s w v k r h r s s f r n r

541  tccattgctg aagggtctca ttcttcaaag aaataaactc agtgacactc ccaagggact
    f p l l k v l i l q r n k l s d t p k g

601  gtggaaactg aagtcattgc agagtttgga totgtcattc aatgggatat tgcaaatagg
    l w k l k s l q s l d l s f n g i l q i

661  gtggtctgat tttcacaact gcctgcaact ggagaatctc tgtttaaaga gcaacaagat
    g w s d f h n c l q l e n l c l k s n k

721  attcaaaatt cccccacaag ctttcaagga cctcaaaaaa ttacaggtca tagaccttag
    i f k i p p q a f k d l k k l q v i d l

781  caacaatgct ctgattacca tcctaccaat gatgatcata gctctagaat ttccccatct
    s n n a l i t i l p m m i i a l e f p h

841  agtggttgac ttggctgata ataactggca gtgtgatgat agtgtggcag tctttcaaaa
    l v v d l a d n n w q c d d s v a v f q

901  ttttatttct gaatcctgga ggaaaaagtg gaatgtcatt tgcaacaggt ctatagggag
    n f i s e s w r k k w n v i c n r s i g

961  tgaggaggcc aacgggggca ctccccagag caggatttcc agggaaaccc gccttcctcc
    s e e a n g g t p q s r i s r e t r l p

1021  cattcatctg catcgcatga aaagcctcat aaggagcaaa gcagagaggc cccagggagg
    p i h l h r m k s l i r s k a e r p q g

1081  aaggcacacg ggcatttcta ctctggggaa gaaggcaaag gccggctctg gtctcaggaa
    g r h t g i s t l g k k a k a g s g l r

```

1141 gaagcagaga cggctgccaa ggagtgttag aagcaccgc gatgtgcagg ctgccggcaa  
 k k q r r l p r s v r s t r d v q a a g  
 ← INSP179-CP2

1201 aaaagaggac gctccccagg acctggctct ggcggtgtgc ctgtcagtgt tcatcacatt  
 k k e d a p q d l a l a v c l s v f i t

1261 ccttgtcgcc ttacgcctgg gggctttcac aaggccttat gttgacagac tgtggcaaaa  
 f l v a f s l g a f t r p y v d r l w q

1321 aaagtgccag agcaaaagcc ctggcctgga caacgcgtat tcaaacgagg gcttctacga  
 k k c q s k s p g l d n a y s n e g f y

1381 tgacatggaa gctgcggggc acacaccaca cccagagacc catctgcgcc aagtatttcc  
 d d m e a a g h t p h p e t h l r q v f

1441 tcatctaagc ctctacgaga accagacccc tttctgggtg acacagccac acccacacgc  
 p h l s l y e n q t p f w v t q p h p h

1501 caccgtaatt cctgatagaa ctctgggaag gagcagaaag gatcctggca gttcgcagag  
 a t v i p d r t l g r s r k d p g s s q

1561 cccaggacag tgcggggaca acaccggggc aggaagtgga aatgatgggtg cagtctattc  
 s p g q c g d n t g a g s g n d g a v y

1621 cattctccag agacatccac atgccggtaa ccgtgaacta atgtcagcag cgcaggacca  
 s i l q r h p h a g n r e l m s a a q d

1681 catccatagg aatgatattc tcggagaatg gacttatgaa actgtggccc aggaagagcc  
 h i h r n d i l g e w t y e t v a q e e

1741 tctcagtga cattcagtgg gcgtctcttc tgtagctggc acgtctcagc ctgtctctgg  
 p l s a h s v g v s s v a g t s h a v s

1801 ctcaagccgt tatgattcca atgaattaga cccttccttc tccggagaaa taacagcttc  
 g s s r y d s n e l d p s l s g e i t a

1861 cctctgtaaa atgctaacac atgcagaagc acagaggact ggagatagta aggaagagg  
 s l c k m l t h a e a q r t g d s k e r

1921 gggcactgaa cagtcacttt gggactcgca gatggaattt tctaaggaaa ggcaagtgag  
 g g t e q s l w d s q m e f s k e r q v

1981 ttcatccatt gatttgctga gcatacagca gccaaaggctg tccggggcaa gggctgagga  
 s s s i d l l s i q q p r l s g a r a e

2041 agcgttttca gccactaca gcgaggttcc atacggtgac ccaagagaca caggcccatc  
 e a l s a h y s e v p y g d p r d t g p

2101 agtctttcct ccaagatggg acagtggcct ggatgtcact cctgctaaca aggaaccagt  
 s v f p p r w d s g l d v t p a n k e p

2161 gcagaaatcc actccttctg acacttgctg tgagttggag agtgactgtg actctgatga  
 v q k s t p s d t c c e l e s d c d s d

2221 ggggtctctg ttcaactctga gctccataag ttcaagagagt gcaaggagca agactgaaga  
 e g s l f t l s s i s s e s a r s k t e

2281 ggcaagtgcct gatgaggagt ccctgcagga cgagagctca ggggcaagca aggacaatgt  
 e a v p d e e s l q d e s s g a s k d n

2341 gacggctgta gacagtcttg aggaaaatgt taccttccaa acaattccag ggaaatgcaa  
v t a v d s l e e n v t f q t i p g k c

2401 gaatcaagaa gatccctttg aaaaacctct catttctgct ccagactctg gcatgtacaa  
k n q e d p f e k p l i s a p d s g m y

2461 gactcatctg gaaaatgcct ctgacactga tagatctgag ggctgtcac cctggcccag  
k t h l e n a s d t d r s e g l s p w p

2521 gtcaccaggg aatagtcctt taggggatga gtttccgggc atgttcactt atgattatga  
r s p g n s p l g d e f p g m f t y d y

2581 cacagctctt caatccaagg cagcagaatg gcattgctca cttagagact tagaattttc  
d t a l q s k a a e w h c s l r d l e f

2641 aaatgtggac gttttacagc aaacaccacc atgttctgct gaagttccct cagatcctga  
s n v d v l q q t p p c s a e v p s d p

2701 taaggctgcc ttccatgaaa gagactcaga cattttaaaa taagaatctt tcattaagga  
d k a a f h e r d s d i l k

2761 aatattcaca gctctaaaca atattctttt aaagatcatg gcaggggaaa actaaagcct  
2821 tcacaacaag atcctgaagg agacaacatg aattccaacc aagtggacac tgatgcaaat  
2881 gaggggttttg tggggccact tgagggcgat gatt

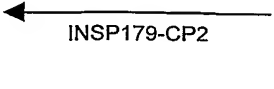
Position and sense of PCR primers →



**Figure 6: INSP179-EC DNA and polypeptide sequence**

1	atgaaaaaac	tctattttcag	agtcattacc	atagttatag	gtcttttattt	tactggaata
	m k n	l y f	r v i t	i v i	g l y	f t g i
	INSP179-CP1					
61	atgacaaatg	catcaagaaa	aagcaatatt	ttatttcaatt	ctgaatgcca	atggaatgaa
	m t n	a s r	k s n i	l f n	s e c	q w n e
121	tatattctga	caaattgttc	ttttaccgga	aagtgtgata	tacctgtgga	catatcacag
	y i l	t n c	s f t g	k c d	i p v	d i s q
181	acagcagcca	ctgtggatgt	aagtttcaat	ttcttttagag	ttctctttaca	gtctcacacg
	t a a	t v d	v s f n	f f r	v l l	q s h t
241	aaaaaagaag	agtggaaaat	aaaacatctg	gacctcagta	acaatctcat	atcaaaaata
	k k e	e w k	i k h l	d l s	n n l	i s k i
301	accttaagcc	cttttgcata	tttacctgct	ttggaagtgt	taaacctcag	caacaatgcc
	t l s	p f a	y l h a	l e v	l n l	s n n a
361	atccactccc	tctcattgga	tctactcagt	cctaagtcct	catgggtgaa	acgccacaga
	i h s	l s l	d l l s	p k s	s w v	k r h r
421	agcagcttca	gaaacaggtt	tccattgctg	aagggtgctca	ttctttcaaag	aaataaaactc
	s s f	r n r	f p l l	k v l	i l q	r n k l
481	agtgacactc	ccaagggact	gtggaaactg	aagtcattgc	agagtttgga	tctgtcattc
	s d t	p k g	l w k l	k s l	q s l	d l s f
541	aatgggatat	tgcaaataag	gtgggtctgat	tttcacaact	gcctgcaact	ggagaatctc
	n g i	l q i	g w s d	f h n	c l q	l e n l
601	tgttttaaaga	gcaacaagat	attcaaaaatt	cccccaacaag	ccttcaagga	cctcaaaaaa
	c l k	s n k	i f k i	p p q	a f k	d l k k
661	ttacagggtca	tagaccttag	caacaatgct	ctgattacca	tcctaccaat	gatgatcata
	l q v	i d l	s n n a	l i t	i l p	m m i i
721	gctctagaat	ttccccatct	agtggttgac	ttggctgata	ataactggca	gtgtgatgat
	a l e	f p h	l v v d	l a d	n n w	q c d d
781	agtggtggcag	tcttttcaaaa	ttttattttct	gaatcctgga	ggaaaaagtg	gaatgtcatt
	s v a	v f q	n f i s	e s w	r k k	w n v i
841	tgcaacaggt	ctataggagg	tgaggaggcc	aacggggggca	ctccccagag	caggatttcc
	c n r	s i g	s e e a	n g g	t p q	s r i s
901	agggaaaacc	gccttcctcc	cattcatctg	catcgcatga	aaagcctcat	aaggagcaaa
	r e t	r l p	p i h l	h r m	k s l	i r s k
961	gcagagaggg	cccaggggagg	aaggcacacg	ggcattttcta	ctctgggggaa	gaaggc aaag
	a e r	p q g	g r h t	g i s	t l g	k k a k

1021 gccggctctg gtctcaggaa gaagcagaga cggctgccaa ggagtgttag aagcaccgc  
a g s g l r k k q r r l p r s v r s t r  
1081 gatgtgcag  
d v q



Position and sense of PCR primers →

**Figure 7: INSP179-EC-SV1 DNA and polypeptide sequence**

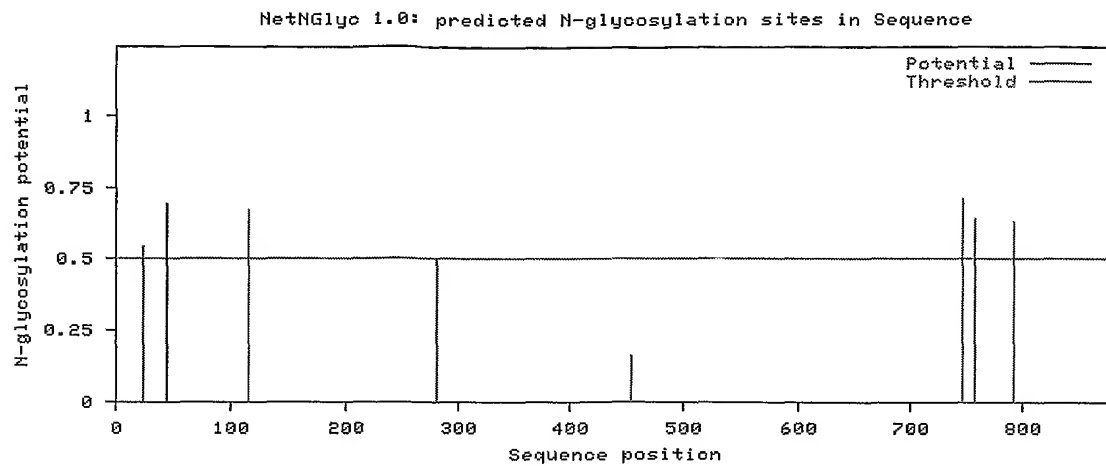
1	atgaaaaaac	tctatttcag	agtcattacc	atagttatag	gtcttttattt	tactggaata
	m k n	l y f	r v i	t i v i	g l y	f t g i
	INSP179-CP1					
61	atgacaaatg	catcaagaaa	aagcaatatt	ttatttcaatt	ctgaatgcca	atggaatgaa
	m t n	a s r	k s n i	l f n	s e c	q w n e
121	tatatctga	caaattgttc	ttttaccgga	aagtgtgata	tacctgtgga	catatcacag
	y i l	t n c	s f t g	k c d	i p v	d i s q
181	acagcagcca	ctgtggatgt	aagtttcaat	ttcttttagag	ttctcttaca	gtctcacacg
	t a a	t v d	v s f n	f f r	v l l	q s h t
241	aaaaaagaag	agtggaaaat	aaaacatctg	gacctcagta	acaatctcat	atcaaaaata
	k k e	e w k	i k h l	d l s	n n l	i s k i
301	accttaagcc	cttttgcata	tttaccatgct	ttggaagtgt	taaacctcag	caacaatgcc
	t l s	p f a	y l h a	l e v	l n l	s n n a
361	atccaactccc	tctcattgga	tctactcagt	cctaagtcct	catgggtgaa	acgccacaga
	i h s	l s l	d l l s	p k s	s w v	k r h r
421	agcagcttca	gaaacagggt	tccattgctg	aagggtgctca	ttcttcaaag	aaataaactc
	s s f	r n r	f p l l	k v l	i l q	r n k l
481	agtgacactc	ccaaggggag	tgaggaggcc	aacgggggca	ctccccagag	caggatttcc
	s d t	p k g	s e e a	n g g	t p q	s r i s
541	agggaaaccc	gccttctctc	cattcatctg	catcgcatga	aaagcctcat	aaggagcaaa
	r e t	r l p	p i h l	h r m	k s l	i r s k
601	gcagagaggg	cccagggagg	aaggcacacg	ggcattttcta	ctctgggggaa	gaaggcaaag
	a e r	p q g	g r h t	g i s	t l g	k k a k
661	gccggctctg	gtctcaggaa	gaagcagaga	cggctgccaa	ggagtgttag	aagcaccgcg
	a g s	g l r	k k q r	r l p	r s v	r s t r
				INSP179-CP2		
721	gatgtgcag					
	d v q					

**Figure 8: Alignment of INSP179-EC and INSP179-EC-SV1 nucleotide sequences**

179EC-SV1	ATGAAAAACCTCTATTTTCAGAGTCATTACCATAGTTATAGGTCTTTATTTTACTGGAATA
INSP179EC	ATGAAAAACCTCTATTTTCAGAGTCATTACCATAGTTATAGGTCTTTATTTTACTGGAATA
179EC-SV1	ATGACAAATGCATCAAGAAAAAGCAATATTTTATTCAATTCTGAATGCCAATGGAATGAA
INSP179EC	ATGACAAATGCATCAAGAAAAAGCAATATTTTATTCAATTCTGAATGCCAATGGAATGAA
179EC-SV1	TATATTCTGACAAATTGTTCTTTTACCGGAAAGTGTGATATACCTGTGGACATATCACAG
INSP179EC	TATATTCTGACAAATTGTTCTTTTACCGGAAAGTGTGATATACCTGTGGACATATCACAG
179EC-SV1	ACAGCAGCCACTGTGGATGTAAGTTTCAATTTCTTTAGAGTTCTCTTACAGTCTCACACG
INSP179EC	ACAGCAGCCACTGTGGATGTAAGTTTCAATTTCTTTAGAGTTCTCTTACAGTCTCACACG
179EC-SV1	AAAAAAGAAGAGTGGAAAAATAAACATCTGGACCTCAGTAACAATCTCATATCAAAAAATA
INSP179EC	AAAAAAGAAGAGTGGAAAAATAAACATCTGGACCTCAGTAACAATCTCATATCAAAAAATA
179EC-SV1	ACCTTAAGCCCTTTTGCATATTTACATGCTTTTGAAGTGTTAAACCTCAGCAACAATGCC
INSP179EC	ACCTTAAGCCCTTTTGCATATTTACATGCTTTTGAAGTGTTAAACCTCAGCAACAATGCC
179EC-SV1	ATCCACTCCCTCTCATTGGATCTACTCAGTCCTAAGTCCTCATGGGTGAAACGCCACAGA
INSP179EC	ATCCACTCCCTCTCATTGGATCTACTCAGTCCTAAGTCCTCATGGGTGAAACGCCACAGA
179EC-SV1	AGCAGCTTCAGAAACAGGTTTCCATTGCTGAAGGTGCTCATTCTTCAAAGAAATAAACTC
INSP179EC	AGCAGCTTCAGAAACAGGTTTCCATTGCTGAAGGTGCTCATTCTTCAAAGAAATAAACTC
179EC-SV1	AGTGACACTCCCAAGGG-----
INSP179EC	AGTGACACTCCCAAGGGACTGTGGAACTGAAGTCATTGCAGAGTTTGGATCTGTCATTC
179EC-SV1	-----
INSP179EC	AATGGGATATTGCAAATAGGGTGGTCTGATTTTCACAACTGCCTGCAACTGGAGAATCTC
179EC-SV1	-----
INSP179EC	TGTTTAAAGAGCAACAAGATATTCAAATTTCCCCACAAGCCTTCAAGGACCTCAAAAAA
179EC-SV1	-----
INSP179EC	TTACAGGTCATAGACCTTAGCAACAATGCTCTGATTACCATCCTACCAATGATGATCATA
179EC-SV1	-----
INSP179EC	GCTCTAGAATTTCCCATCTAGTGGTTGACTTGGCTGATAATAACTGGCAGTGTGATGAT
179EC-SV1	-----
INSP179EC	AGTGTGGCAGTCTTTCAAATTTTATTTCTGAATCCTGGAGGAAAAAGTGAATGTCATT
179EC-SV1	-----
INSP179EC	-----GAGTGAGGAGGCCAACGGGGGCACTCCCCAGAGCAGGATTTCC
179EC-SV1	AGGGAAACCCGCCTTCTCTCCATTCTGTCATCGCATGAAAAGCCTCATAAGGAGCAAA
INSP179EC	AGGGAAACCCGCCTTCTCTCCATTCTGTCATCGCATGAAAAGCCTCATAAGGAGCAAA
179EC-SV1	GCAGAGAGGCCCCAGGGAGGAAGGCACACGGGCATTTCTACTCTGGGGAAGAAGGCAAAG
INSP179EC	GCAGAGAGGCCCCAGGGAGGAAGGCACACGGGCATTTCTACTCTGGGGAAGAAGGCAAAG
179EC-SV1	GCCGGCTCTGGTCTCAGGAAGAAGCAGAGACGGCTGCCAAGGAGTGTTAGAAGCACCCGC
INSP179EC	GCCGGCTCTGGTCTCAGGAAGAAGCAGAGACGGCTGCCAAGGAGTGTTAGAAGCACCCGC
179EC-SV1	GATGTGCAG
INSP179EC	GATGTGCAG

**Figure 9: Alignment of predicted INSP179 and the cloned INSP179-EC and INSP179-EC-SV1 polypeptide sequences**

INSP179-EC_PREDICTION	MKNLYFRVITIVIGLYFTGIMTNASRKSNIILFNSECQWNEYILTNCSEFTG
INSP179-EC_CLONED	MKNLYFRVITIVIGLYFTGIMTNASRKSNIILFNSECQWNEYILTNCSEFTG
INSP179-EC-SV1_CLONED	MKNLYFRVITIVIGLYFTGIMTNASRKSNIILFNSECQWNEYILTNCSEFTG
	*****
INSP179-EC_PREDICTION	KCDIPVDISQTAATVDVSNFFRVLLQSHTKKEEWKIKHLDLSNNLISKI
INSP179-EC_CLONED	KCDIPVDISQTAATVDVSNFFRVLLQSHTKKEEWKIKHLDLSNNLISKI
INSP179-EC-SV1_CLONED	KCDIPVDISQTAATVDVSNFFRVLLQSHTKKEEWKIKHLDLSNNLISKI
	*****
INSP179-EC_PREDICTION	TLSPFAYLHALEVLNLSNNAIHSLSLDLLSPKSSWVKRHRSSFRNRFPLL
INSP179-EC_CLONED	TLSPFAYLHALEVLNLSNNAIHSLSLDLLSPKSSWVKRHRSSFRNRFPLL
INSP179-EC-SV1_CLONED	TLSPFAYLHALEVLNLSNNAIHSLSLDLLSPKSSWVKRHRSSFRNRFPLL
	*****
INSP179-EC_PREDICTION	KVLILQRNKLSDTPKGLWKLKSLQSLDLSFNGILQIGWSDFNCLQLENL
INSP179-EC_CLONED	KVLILQRNKLSDTPKGLWKLKSLQSLDLSFNGILQIGWSDFNCLQLENL
INSP179-EC-SV1_CLONED	KVLILQRNKLSDTPK-----
	*****
INSP179-EC_PREDICTION	CLKSNKIFKIPPOAFKDLKKLQVIDLSNNALITILPMMIIALEFPHLVVD
INSP179-EC_CLONED	CLKSNKIFKIPPOAFKDLKKLQVIDLSNNALITILPMMIIALEFPHLVVD
INSP179-EC-SV1_CLONED	-----
INSP179-EC_PREDICTION	LADNNWQCDDSVAVFQNFISESWRKKWNVICNRSIGSEEANGGTPQSRIS
INSP179-EC_CLONED	LADNNWQCDDSVAVFQNFISESWRKKWNVICNRSIGSEEANGGTPQSRIS
INSP179-EC-SV1_CLONED	-----GSEEANGGTPQSRIS
	*****
INSP179-EC_PREDICTION	RETRLPPPIHLHRMKSILIRSKAERPQGGRHTGISTLGKKAKAGSGLRKKQR
INSP179-EC_CLONED	RETRLPPPIHLHRMKSILIRSKAERPQGGRHTGISTLGKKAKAGSGLRKKQR
INSP179-EC-SV1_CLONED	RETRLPPPIHLHRMKSILIRSKAERPQGGRHTGISTLGKKAKAGSGLRKKQR
	*****
INSP179-EC_PREDICTION	RLPRSVRSTRDVQ
INSP179-EC_CLONED	RLPRSVRSTRDVQ
INSP179-EC-SV1_CLONED	RLPRSVRSTRDVQ
	*****

**Figure 10: Predicted INSP179 glycosylation sites**

SeqName	Position	Potential	Score
Sequence	23 NASR	0.5420	(6/9)
Sequence	45 NCSF	0.6932	(9/9)
Sequence	115 NLSN	0.6702	(9/9)
Sequence	282 NRSI	0.4953	(4/9)
Sequence	454 NQTP	0.1590	(9/9)
Sequence	746 NVT A	0.7102	(9/9)
Sequence	756 NVTF	0.6390	(9/9)
Sequence	792 NASD	0.6288	(6/9)

